[redacted]

Western Management Science Institute
University of California • Los Angeles

University of California

Los Angeles

Western Management Science Institute

Working Paper No. 77

THE ECONOMICS OF UNCERTAINTY XIII

by

Karl Borch

May 1965

This paper consists of notes prepared for a series of lectures given at the Graduate School of Business Administration at the University of California, Los Angeles, during the academic year 1964-65. The work has been supported by the Western Management Science Institute under a grant from the Ford Foundation, and by the Office of Naval Research under Task 047-041. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Chapter XIII

Credibility and Subjective Probabilities

13.1 In the preceding chapters we have generally assumed that the decision maker knew the probabilities with which the different states of the world would occur. In practice a decision maker will often argue that he does not really know these probabilities, although he is not so "completely ignorant" that he feels he should use the Laplace principle of insufficient reason, i.e. assign equal probabilities to all states of the world.

In this Chapter we shall study the ways in which such vague knowledge or beliefs can be brought to bear on the decision problem. The ideas which we shall develop were first explicitly formulated by Savage [5].

To make our discussion concrete, we shall consider an insurance company, which holds a portfolio of insurance contracts, and reserve funds amounting to S. Let total claims payable under the contracts in the portfolio be a stochastic variable with the distribution $G(x)$. We shall assume that this distribution is known to the company.

The company will assign the following utility to this situation

$$U(S) = \int_0^{\infty} u(S-x)\, dG(x)$$

where $u(x)$ is the utility function which represents the company's preference ordering. Since we have not brought the time element into the

model, we shall assume that all contracts in the portfolio expire within a fairly short period.

13.2. Let us now assume that the company is offered an additional contract which will expire within the same period.

Let P = the premium paid for this contract, and let the claim distribution be

$$\begin{aligned} &Z \text{ with probability } p \\ &0 \text{ with probability } 1-p=q \end{aligned}$$

The company will accept the new contract if, and only if it leads to an increase in utility, i.e. if

$$pU(S+P-Z) + q\,U(S+P) > U(S)$$

In some contexts it may be more natural to assume that company is invited to "quote a premium" for the new contract. The equation

$$pU(S+P-Z) + qU(S+P) = U(S)$$

will then determine the lowest premium P which the company can quote.

This is a very simple application of the principles, which we have developed in earlier chapters.

However, in a real life situation, the company may not feel so certain that the relevant probability is exactly p.

Our problem is to find out what this actually may mean, and to study how the insurance company will, or should, make its decision in this situation.

13.3 Let us first assume that the insurance company maintains that it <u>knows absolutely nothing</u> about the risk covered by the new contract. If this statement has any meaning at all, it must imply that any value of p between 0 and 1 is equally possible - or equally probable.

It is then natural to write the equation from the preceding paragraph in the following form

$$p\{ U(S+P-Z) - U(S+P) \} + U(S+P) = U(S)$$

and multiply by dp, and integrate from 0 to 1. This will give

$$\frac{1}{2} \{ U(S+P-Z) + U(S+P) \} = U(S)$$

as the equation, which determines the lowest premium P which the company can quote.

13.4 In practice we will not often have to make decisions under complete ignorance. We will usually have some information or <u>prior belief</u> about p. The mere fact that somebody wants to pay for this insurance contract, indicates that p is not zero - i.e. the event which will lead to a claim payment is not impossible.

The usual procedure may be as follows: The actuary and the more or less experienced underwriters, may agree that the "best estimate" is, say $p = 0.10$, adding that this is little more than an educated guess. When pressed for more precision, they may state that p is very likely to be somewhere in the interval $(0.05, 0.20)$, or that they are certain that $p < 0.40$.

Arguments of this kind reflect a vague, but very real feeling of uncertainty. The most natural way of giving precision to this statement seems to be to specify the weights which should be given to the various possible values of p. We can do this by specifying a function $f(p)$, which takes its greatest value for the "best estimate" or "most likely" value of p.

There is obviously nothing to prevent us from normalizing the function, and requiring that

$$\int_0^1 f(p)\, dp = 1$$

This means that $f(p)$ can be interpreted as a probability distribution, which represents our belief about the value of p.

So far this makes sense. The trouble comes if or when we state something like

$f(0.1) = \Pr\{ p = 0.1 \} =$ the probability that the parameter p is equal to 0.1.

This statement has no real meaning. A parameter is not a stochastic variable, so it has no meaning to assign a probability other than 0 and 1 to the "event" that it takes a particular value.

13.5 If our insurance company can specify the function which represents its prior belief, the decision problem is solved by multiplying the basic equation by $f(p)\, dp$ and integrating from 0 to 1. This gives

$$U(S+P) + \{ U(S+P-Z) - U(S+P) \} \int_0^1 pf(p)\, dp = U(S)$$

or if we write

$$\bar{p} = \int_0^1 pf(p)\, dp$$

$$\bar{p}\, U(S+P-Z) + (1 - \bar{p})\, U(S+P) = U(S)$$

This means that the company acts as if it was certain the parameter has the value $\bar{p}$

13.6. The example we have discussed, is very artificial, but it brings out the essential idea involved.

As a more realistic example, let us assume that the company is offered a portfolio of n contracts of the type considered in the example. In this case the expected claim payment is $npZ = \bar{y}$ and the amount of premium received is nP.

The "Principle of Equivalence" which is the foundation of classical insurance theory, requires that expected payments and receipts shall be equal, i.e. that the premium for each of these contracts shall be

$$P = pZ$$

This will, however, not be acceptable to a company which has a "risk aversion", i.e. a company which is worried about the possibility that actual payments may exceed the expected value.

It is easy to see that

$$Pr\,\{y = kZ\} = \binom{n}{k} p^k (1 - p)^{n-k}$$

Hence the minimum acceptable premium is determined by

$$U(S) = \sum_{k=0}^{n} u\,(S + nP - kZ) \binom{n}{k} p^k (1-p)^{n-k}$$

This formula takes into account the uncertainty which in statistical language

is due to "sampling fluctuations".

The uncertainty, due to incomplete knowledge about the true value of the parameter, is logically of a different nature. If, however, we are willing to specify a <u>subjective probability</u> distribution $f(p)$, which represents our "prior belief", we can deal with this second kind of uncertainty in the classical way.

13.7 In general the problem is formulated as follows:

Claim payment under a portfolio of insurance contracts, is a stochastic variable with a distribution $G(x,\alpha)$, or we can take α to be a vector with mean, variance and other parameters of the distribution as elements. For the sake of simplicity, we shall assume that the distribution is continuous, and that

$$g(x,\alpha) = \frac{\partial G(x,d)}{\partial x}$$

If the company has to quote a premium for this portfolio, it will compute the expected utility

$$\int_0^\infty U(S+P-x)\, g(x,\alpha)\, dx$$

If there is some further uncertainty about the parameter α, expressed by a prior distribution $f(\alpha)$, we have to carry out another integration

$$\int_A \{ \int_0^\infty U(S+P-x)\, g(x,\alpha)\, dx \}\, f(\alpha)\, d\alpha$$

where A is the domain of α.

However, the latter integral is obviously equal to

$$\int_0^\infty U(S+P-x) \left(\int_A g(x, \alpha) f(\alpha) d\alpha \right) dx$$

The inner integral

$$h(x) = \int_A g(x, \alpha) f(\alpha) d\alpha$$

can of course be interpreted as a probability distribution.

This means that the lowest acceptable premium P is determined by

$$\int_0^\infty U(S+P-x) h(x) dx = U(S)$$

This is a solution of the same form as the one we found earlier. The whole elaborate reasoning about uncertainty over the value of the parameters, means only that we replace the original distribution $g(x,\alpha)$ by $h(x)$.

13.8 The procedure of the preceding paragraph has some practical value only if we know - or have good reasons to believe - that claim payments really are generated by a distribution of the form $g(x,\alpha)$.

In practice we do not often know this. In insurance one will usually analyse the amounts paid as claim compensation under a large number of identical contracts. Let us, for instance, assume that in a portfolio of 5000 automobile collision insurance contracts we have observed 380 claims, leading to a total payment of \$ 350,000.

Table 1

Claim Payment	Number of Claims
0	4620
$ 0-100	0
$ 101-500	100
$ 501-1000	105
$ 1001-1500	110
$ 1501-2000	25
$ 2001-2500	30
$ 2501-3000	4
$ 3000-3500	6

Let us further assume that claim payments can be broken down in more detail as shown in Table 1.

In this situation we can ignore the detailed breakdown, and just note that the average claim payment per contract is $70 and on this basis formulate our beliefs about claim payments in the next portfolio which our company will underwrite.

A more "sophisticated" approach may be to fit a distribution to the data of the table - for instance by the "method of moments" - and indicate the reliability of the estimated parameter. In doing so we may, however, well have added new <u>assumptions</u> rather than new <u>knowledge</u> to the model, and it is quite possible that the simpler approach may be the sounder policy

13.9 Let us now return to our simple example, in which the claim payment could take only the values 0 or Z. The only unknown parameter

was p = the probability that the claim should become payable.

If we really have no idea about the true value of p, there may be something to be said for working on the assumption that all values of p should be given the same weight. Such cases are, however, rare, so let us therefore assume that we have some relevant knowledge, namely that in a portfolio of n comparable contracts, there were k claim payments.

To a statistician, it is then natural to suggest that we act as if

$$p = \frac{k}{n}$$

He will usually be able to justify this in several different ways.

There is, however, some uncertainty about this estimate, particularly if n is small, and we want to allow for this in our decision.

Let us therefore write:

$$Pr(k|p) = \binom{n}{k} p^k (1-p)^{n-k}$$

This is the probability of the observed result k, if the true probability is p. This is usually referred as the <u>likelihood</u> of the observed result. One justification for taking $p = k/n$ is that this value will maximize the likelihood.

13.10 From the theory of conditional probability we know that

$$Pr(k|p)\ Pr(p) = Pr(p|k)\ Pr(k)$$

or

$$Pr(p|k) = \frac{Pr(k|p)\ Pr(p)}{Pr(k)}$$

Here the denominator can be interpreted as the absolute probability of k - i.e. the probability of observing k-claims, regardless of what the true probability p may be.

We then have in a purely formal way

$$Pr(k) = \Sigma_p \; Pr(k|p) \; Pr(p)$$

Here the sum is over all values of p, and Pr(p) is the weight of our belief that the true value of the parameter is p. To express this result in the notation we have used earlier, we shall write

$$Pr(p) = f(p)$$

Hence our formula can be written

$$Pr(p|k) = \frac{\binom{n}{k} p^k (1-p)^{n-k} \; f(p)}{\int_0^1 \binom{n}{k} \; p^k (1-p)^{n-k} \; f(p) \; dp}$$

This is a special case of the classical Bayes' formula.

The formula gives us the "likelihood" that p is the true parameter, given that k was observed. The formula depends on the "prior belief", represented by the density function f(p).

$Pr(p|k)$ can therefore be taken to be a distribution, representing our belief about p, when we combine the statistical experience and our prior belief.

13.11 If we know nothing about p, except that k claims occurred in a sample of n, it may seem natural to assume $f(p) = 1$. This will reduce our formula to

$$Pr\,(p|k) = \frac{p^k (1-p)^{n-k}}{\int_0^1 p^k (1-p)^{n-k} \; dp}$$

The denominator in this expression is the so-called Beta-function

$$\int_0^1 p^k(1-p)^{n-k}\,dp = B(k+1,\ n-k+1) = \frac{k!\,(n-k)!}{(n+1)!}$$

We can now apply this result to our original problem. We started in para 13.3 with the equation

$$U(S+P) - p\ \{\ U(S+P) - U(S+P-Z)\ \} = U(S)$$

We then multiplied by $f(p)\,dp$ and integrated over p from 0 to 1, and obtained

$$U(S+P) - \bar{p}\ \{U(S+P) - U(S+P-Z)\ \} = U(S)$$

where

$$\bar{p} = \int_0^1 p\ f(p)\ dp$$

In our present example we have to replace $f(p)$ by

$$Pr(p|k) = \frac{1}{B(k+1,\ n-k+1)}\ p^k(1-p)^{n-k}$$

Substituting this, we obtain

$$\bar{p} = \frac{k+1}{n+2}$$

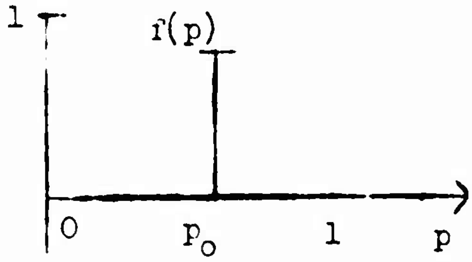
This appears as the probability which we should use in our decision, if we want to combine the experience obtained by observing a comparable portfolio, and our prior belief that every value of p was equally likely.

We see that for $k = 0$, we should take $\bar{p} = \frac{1}{n+2}$. This means that unless n is very large, our prior belief will still carry some weight. The fact that no claim occurred in a sample of n insurance contracts, does not lead us to make future decisions on the assumption that $\bar{p} = 0$.

13.12 So far we have treated prior belief as a very vague concept - even vaguer than the utility concept. Our only concrete result was that if prior belief shall make any sense, it must be possible to represent it as a probability distribution - or a weight function - over the set of values which can be taken by some parameter.

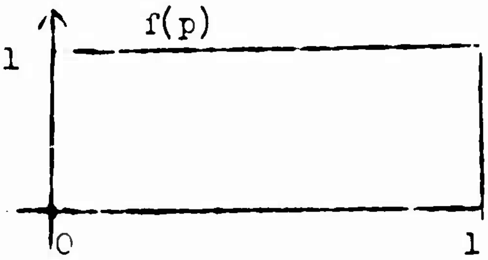
If we are absolutely certain that this parameter has the value p_o, the distribution will be as in Fig. 1.

Fig. 1



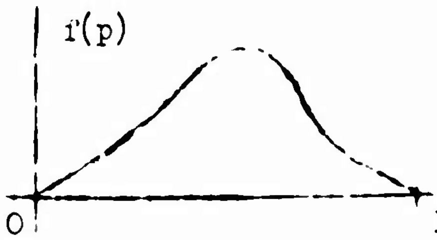
If we think any value between 0 and 1 as equally likely, the distribution will be as in Fig. 2.

Fig. 2



In the intermediary cases our prior belief will be represented by a function as illustrated by Fig. 3.

Fig. 3



13.13 From these considerations it follows that we are interested only in the general shape of the prior distribution. We want a distribution function, which can represent prior beliefs with sufficient approximation - or with as much precision as the decision maker can express.

We are therefore led to examine if this class is sufficiently rich to represent all the prior beliefs, which we may want to study.

For the mean of the distribution we find:

$$\mu = \int_0^1 p\, f(p)\, dp = \frac{a+1}{a+b+2}$$

and for the variance:

$$\sigma^2 = \int_0^1 (p-\mu)^2\, f(p)\, dp = \frac{(a+1)\,(b+1)}{(a+b+2)^2\,(a+b+3)}$$

From these expressions we see that if μ and σ are given, we can usually determine a and b. This means that if we describe our beliefs by specifying only the two first moments of the prior distribution, we can always find a Beta-distribution which meets our specifications.

13.15 Let us now assume that our prior belief can be represented by the distribution:

$$f(p) = \frac{1}{B(a+1,\ b+1)} \quad p^a(1-p)^b$$

Let us assume that when we in this way have made up our mind as to what we believe about the claim frequency, we learn that in a comparable portfolio of n contracts there were k claims. To make use of this new knowledge, we apply the formula in para 13.10 and find:

$$Pr(p|k) = \frac{p^{k+a}(1-p)^{n-k+b}}{\int_0^1 p^{k+a}(1-p)^{n-k+b}} = \frac{p^{k+a}(1-p^{n-k+1})}{B(k+a+1,n-k+b+1)}$$

This gives the probability to be used for our decision:

$$\bar{p} = \int_0^1 p\, Pr(p|k)\, dp = \frac{k+a+1}{n+a+b+1}$$

If a prior belief, we feel certain that $p = p_0$, we must have:

$$p_0 = \mu = \frac{a+1}{a+b+2} > 0$$

and

$$\sigma^2 = \frac{(a+1)(b+1)}{(a+b+2)^2(a+b+3)} = 0$$

It is obvious that in this case both a and b must be infinite.

From the first condition we obtain

$$p^0 = \frac{\frac{a}{b} + \frac{1}{b}}{\frac{a}{b} + 1 + \frac{2}{b}}$$

or

$$\lim \frac{a}{b} = \frac{p_0}{1-p_0}$$

From the expression for $\bar{p}$ we find

$$\bar{p} = \frac{\frac{k}{b} + \frac{a}{b} + \frac{1}{b}}{\frac{n}{b} + \frac{a}{b} + 1 + \frac{1}{b}}$$

Going to the limit, we find

$$\bar{p} = \frac{\frac{p_0}{1 - p_0}}{\frac{p_0}{1-p_0} + 1} = p_0$$

This expresses the obvious. If we are certain that $p = p_0$, we will make our decision accordingly, no matter what experimental evidence should become available.

13.16 If in this example our prior belief can be represented by a Beta-distribution

$$f(p) = \frac{1}{B(a+1,\ b+1)}\ p^{a}(1-p)^{b}$$

can give this a simple intuitive meaning:

Our beliefs about the parameter p are as if we had observed that a claims were made in a portfolio of $a+b$ insurance contracts.

If we actually had made this observation, we would have some knowledge or belief about p. Our problem is then to express formally what this knowledge really is. The natural way of doing this - at least to a statistician - is to say that our knowledge is represented by a Beta - distribution over the domain of the possible values of the parameter.

However we really want to carry the argument through in the opposite direction. We want to start with the prior beliefs which we have about the parameter p and give a precise description of these beliefs. We can do this by specifying a "prior distribution", but it would be more attractive if we could describe an experiment and a particular outcome of the experiment which in some sense represents our beliefs about the parameter.

13.17 The problem we outlined in the preceding paragraph has been discussed for more than 50 years - often in an obscure language - by American actuaries under the name of credibility theory. This theory was founded by Whitney [6] and has been developed by Perryman [3], Bailey [1], Carlson [2], and others, without much contact with the mainstream of statistical theory.

To illustrate the application of the theory, let us consider an insurance company which has to quote a premium for an insurance contract of the simple type considered in our previous examples. Let us assume that the company

when making this decision can draw on two types of information:

(i) Statistical information about comparable contracts - for instance that there were k claims in a portfolio of n identical or similar contracts.

(ii) Other relevant information - for instance statistical observations of portfolios of contracts which are not quite comparable to the contract in question.

If there is sufficient statistical information, i.e. if n is large, the company will not consider the other information. The company will act as if it was certain that $p = k/n$. In this case the actuaries will say that the statistical experience carries "100 per cent credibility", and they will usually be embarrassed if they are asked to justify this statement.

When the statistical experience is insufficient, the company may use other relevant information. However, this is not possible unless different pieces of information can be made comensurable. This leads us to determine the statistical evidence which is equivalent to the other relevant information which we want to use, and this is just what we did in para 13.15.

Theoretically an insurance company should bring into consideration additional information until the equivalent statistical experience carries 100 per cent credibility. How this should be done is a difficult problem, which is far from being satisfactorily solved in the existing theory. The statistical experience of an insurance company which has written 1 million fire-insurance contracts in New York State obviously contains information which may be of value to a company which writes fire insurance in California.

Intuitively we may feel that this information is of less value than the information we could have obtained from the statistical experience of a similar company operating in California. It is, however, not easy to give a precise formulation to such feelings. Will 1 million observations from New York be equivalent to 800,000 observations from California?

If an insurance company operates with a system of premium rates derived from statistical experience which carries 100 per cent credibility, good or bad underwriting results will be explained as caused by random fluctuations. These results will not induce the company to change its premium rates.

In practice an insurance company will not usually assign 100 per cent credibility to the statistical experience which constitutes the foundation of its premium rates. The most obvious reason for this cautious attitude is that the basic probabilities may change with time.

In this situation the company will accumulate new statistical experience as time goes by, and this new information may lead the company to adjust its premium rates. How much the rates should be changed will depend on the credibility carried by the initial statistical experience. This question can be the subject of heated discussions between company representatives and State Insurance Commissioners.

13.18 The Beta-distribution is not always a convenient representation of our "prior belief". Let us, for instance, assume that we are considering an investment, where we know that the return is normally distributed with unit variance and mean m, i.e. in our notation we have:

$$g(x,m) = \frac{1}{\sqrt{2\pi}} \; e^{-\frac{1}{2}(x-m)^2}$$

If our prior belief about m can be represented by the distribution $f(m)$, we should then compute

$$\bar{g}(x) = \frac{1}{2\pi} \int_a^b e^{-\frac{1}{2}(x-m)} f(m)\, dm$$

and use this distribution $\bar{g}(x)$ to compute the expected utility, which will be our decision criterion.

If we in this integral take $f(m)$ as a Beta-distribution - adjusted so that it applies to the interval (a,b), we will get into some very messy computations, and we will have to work with an extremely inconvenient function $\bar{g}(x)$.

The natural conjugate distribution in this case is

$$f(m) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(m-\mu)^2}{2\sigma^2}}$$

This will give

$$\bar{g}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(x-m)^2} e^{-\frac{(m-\mu)^2}{2\sigma^2}} dm$$

or

$$\bar{g}(x) = \frac{1}{\sqrt{2(1+\sigma^2)}} e^{-\frac{1}{2(1+\sigma^2)}(x-\mu)^2}$$

This means that the function, which we use in our decision problem - the so-called <u>posterior distribution</u> - is normal. We may note that the variance of this distribution is the sum of the variances from the two distributions we started with. The choice of the appropriate conjugate distribution has been discussed in great detail by Raiffa and Schlaifer [4].

13.19 If we draw a sample of n from a Normal population with unit variance, the stochastic variable

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + \dots x_n)$$

will be Normally distributed with variance $\frac{1}{n}$.

If the population mean is μ the probability that the average of our sample shall be equal to $\bar{x}$, is proportional to

$$e^{-\frac{n}{2}(\bar{x} - \mu)^2}$$

The best estimate of the population mean is then the value of μ, which maximizes this likelihood function, i.e. $\mu = \bar{x}$.

This means that if we represent our prior belief about m by the distribution

$$\frac{1}{\sigma\sqrt{2\pi}^2} \quad e^{-\frac{(m-\mu)^2}{2\sigma^2}}$$

we feel as confident about $m = \mu$ as we would be if μ was the observed average of a sample of $n = 1/\sigma$

If in this example our prior belief cannot be represented by a Normal distribution - for instance because our beliefs are not symmetrical round some central value μ, we will have to go through more cumbersome mathematics. We will also have to give up the intuitively attractive idea of stating that our beliefs are equivalent to the beliefs we would have if a specific experiment had given a particular outcome.

13.20 Let us now return to the results of para 13.7. We found that our final decision would be based on the expected utility

$$\int_0^\infty \left(\int_A u(S + P - x)\, g(x,\alpha)\, f(\alpha)\, d\alpha \right) dx$$

In this expression:

(i) The prior distribution $f(\alpha)$ represents what we believe

(ii) The utility function $u(x)$ represents what we want

(iii) The distribution $g(x,\alpha)$ represents what we know.

All these three elements must be considered in a rational decision, and in an analysis of the problem they should be separated.

In practice it may, however, be difficult to separate what we believe from what we know. In our simple example the place of the distribution $g(x,\alpha)$ was taken by the binomial

$$\binom{n}{k} p^k (1-p)^{n-k}$$

where p is the unknown parameter. This binomial distribution rests on the assumptions:

(i) The probability of a claim is the same under all the n contracts

(ii) The probability of a claim under an arbitrary contract is independent of whether claims have been made under any of the $n-1$ other contracts

If we <u>know</u> that these assumptions are true, there is no problem. If, however, these assumptions just represent our beliefs, or are accepted as working hypotheses, they should be included in $f(\alpha)$ and not in $g(x,\alpha)$. This means that the separation of the different elements, which seems essential to a rational analysis of the decision problem, is by its very nature arbitrary. This again means that in a preliminary study only the general shape of the functions f, g, and u is significant, and that we should feel free to choose functions which are easy to manipulate mathematically.

REFERENCES

[1] BAILEY, A. L. : "A Generalized Theory of Credibility", Proceedings of the Casualty Actuarial Society XXXII (1945)

[2] CARLSON, T. O. "Observations on Casualty Insurance Rate-making Theory in the United States", Transactions of the 17th International Congress of Actuaries Vol III pp 541 - 559, London 1964.

[3] PERRYMAN, F. S. "Some Notes on Credibility", Proceedings of the Casualty Actuarial Society, XIX (1932)

[4] RAIFFA, H. and R. SCHLAIFER: Applied Statistical Decision Theory, Harvard, 1961.

[5] SAVAGE, L. J. : The Foundations of Statistics Wiley 1954

[6] WHITNEY, A. W.: "The Theory of Experience Rating" Proceedings of the Casualty Actuarial Society IV (1918)